BBC RD 1989/8

Research Department Report

Comparison of DPCM prediction strategies for high-quality digital television bit-rate reduction

M.J. Knee, M.A., A.F.I.M.A., N.D. Wells, B.A., D.Phil.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION

BBC RD 1989/8

COMPARISON OF DPCM PREDICTION STRATEGIES FOR HIGH-QUALITY DIGITAL TELEVISION BIT-RATE REDUCTION

M.J. Knee, M.A., A.F.I.M.A. N.D. Wells, B.A., D.Phil.

Summary

This Report compares some DPCM prediction strategies for bit-rate reduction of YUV component-coded digital television signals for broadcast-quality contribution or distribution applications. Several systems are compared for two critical moving picture sequences as a test of worst-case performance. The effects of transmission errors on the decoded picture are also compared for some of the systems. The results show that, in terms of picture quality, many of the systems give similar performance on critical sequences but that the error performance of multi-element predictors is much better than that of some of the adaptive prediction strategies. The best performance is obtained from a least-mean-square (LMS) adaptive predictor.

Issued under the Authority of

M.E.B. Moffat

Head of Research Department

Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION

(EL-200)

1989

© British Broadcasting Corporation

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior permission.

COMPARISON OF DPCM PREDICTION STRATEGIES FOR HIGH-QUALITY DIGITAL TELEVISION BIT-RATE REDUCTION

M.J. Knee, M.A., A.F.I.M.A.   N.D. Wells, B.A., D.Phil.

1. Introduction 1

2. Description of Prediction Strategies 1

   2.1 Category A: Non-adaptive or fixed (linear) predictors 1

   2.2 Category B: Adaptive prediction with transmission of the adaptation control signal 2

   2.3 Category C: Adaptive prediction without explicit transmission of the adaptation control signal 3

   2.4 Category D: Continuously adaptive prediction 3

3. Comparison of Prediction Strategies 4

   3.1 Source pictures 4

   3.2 Techniques for comparison 4

   3.3 Results of comparisons for moving picture sequences 6

   3.4 Results of comparison for stationary picture material 6

   3.5 Comparison between fixed- and variable-length coding 8

4. Measurements of Error Performance 8

5. Conclusions 9

6. Acknowledgements 10

7. References 10

Appendix: Variation of *S/N* with Entropy 10

© BBC 2005. All rights reserved. Except as provided below, no part of this document may be reproduced in any material form (including photocopying or storing it in any medium by electronic means) without the prior written permission of BBC Research & Development except in accordance with the provisions of the (UK) Copyright, Designs and Patents Act 1988.

The BBC grants permission to individuals and organisations to make copies of the entire document (including this copyright notice) for their own internal use. No copies of this document may be published, distributed or made available to third parties whether by paper, electronic or other means without the BBC's prior written permission. Where necessary, third parties should be directed to the relevant page on BBC's website at http://www.bbc.co.uk/rd/pubs/ for a copy of this document.

COMPARISON OF DPCM PREDICTION STRATEGIES FOR HIGH-QUALITY DIGITAL TELEVISION BIT-RATE REDUCTION

M.J. Knee, M.A., A.F.I.M.A. N.D. Wells, B.A., D.Phil.

1. INTRODUCTION

This Report compares some DPCM bit-rate reduction strategies for *YUV* component-coded digital television signals for broadcast-quality contribution or distribution applications. The performance of DPCM systems is determined mainly by the method used to predict the input signal and this Report concentrates, therefore, on a comparison of various prediction strategies.

DPCM is a popular bit-rate reduction technique and many variations of the method have been reported in the literature. Comparative interpretation of the published results is often difficult because of the diversity of test methods, assumptions and source picture material used by the various authors. In order to assist the process of selection of candidate DPCM systems for international standardization, some comparative studies were made of several different DPCM schemes. This Report presents the results of those studies, which concentrate on (a) the coding performance and (b) the behaviour in the presence of transmission errors of the various systems.

The particular application which has been considered is the transmission of high-quality component-coded video signals at between 2.5 and 3 bits per sample for the luminance component. Depending on the number of bits per sample allocated to the chrominance component, this would give a total bit rate for both luminance and chrominance of up to 60 Mbit/s for the transmission of video signals sampled according to CCIR Recommendation 601. With some initial sample rate reduction of the components, the total video bit rate would be up to 30 Mbit/s.

The principle of DPCM coding is illustrated in Fig. 1. The prediction of each input sample value is made using a weighted sum of previous sample values. The difference between the input value and the prediction is nonlinearly quantized and the quantized prediction error is then coded for transmission. Variable-length coding techniques, such as Huffman coding, can be applied to the quantized prediction error and this supplements the bit-rate reduction provided by the nonlinear quantizer.

It is important, whether variable-length coding is used or not, to maximize the accuracy of the prediction in order to minimize quantizing distortion in the decoded picture at the low bit rates being considered. Existing and proposed DPCM systems use very different prediction strategies which, broadly speaking, fall into four categories:

A: Non-adaptive (linear) prediction.

B: Adaptive prediction with transmission of the adaptation control signal (explicit signalling).

C: Adaptive prediction without transmission of the adaptation control signal (implicit signalling).

D: Continuously adaptive prediction (with or without transmission of the adaptation control signal).

Examples of these different prediction strategies are compared. Section 2 gives a brief description of each of the strategies investigated, Section 3 covers the relative quality performance (and includes a comparison with two-dimensional transform coding) and Section 4 gives a comparison of decoder error performance.

2. DESCRIPTION OF PREDICTION STRATEGIES

2.1 Category A: Non-adaptive or fixed (linear) predictors

The prediction, $p(x,y,t)$ is a weighted sum of past decoded sample values, $s(x,y,t)$, the weighting coefficients being fixed:

$$p(x,y,t) = \sum g(i,j,k)\, s(x-i,y-j,t-k)$$

where x, y and t are the horizontal, vertical and temporal coordinate values expressed in samples, *picture* lines and frames respectively and the summation is taken over the set of predictor elements. Note that, in the description of the predictor elements, positive values of i, j or k correspond to sample locations occurring earlier in time.

The following four fixed predictors have been studied.

A1: Typical two-dimensional predictor

A typical two-dimensional predictor was used with the following coefficients: $g(1,0,0) = 0.625$, $g(0,2,0) = 0.125$, $g(-1,2,0) = 0.25$. The performance of other 2-D predictors of similar complexity were examined and found to be similar.

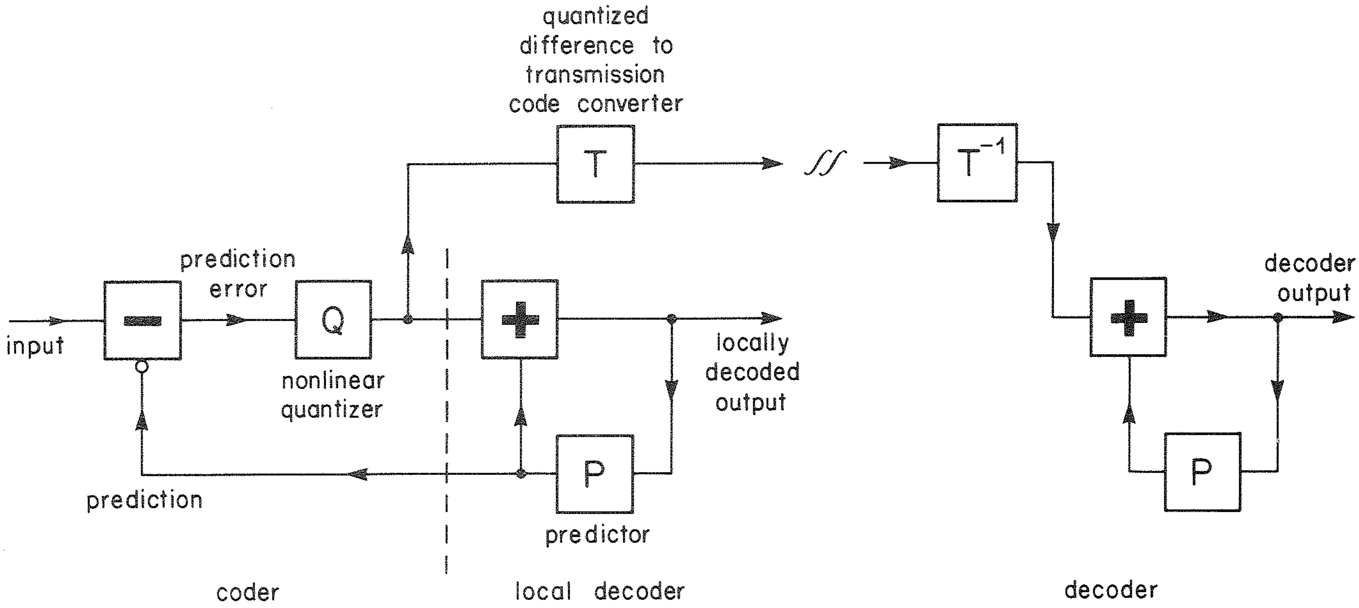


Fig. 1 - Block diagram of DPCM coder and decoder.

A2: Pure interframe predictor

A pure interframe predictor has one element, $g(0,0,1) = 1$. This predictor is excellent for still pictures.

A3a: Two-field multi-element predictor

This type of predictor has been used in DPCM systems developed at the BBC Research Department[1] and its advantages are that it gives good performance, is straightforward to implement and generally gives a system which is very stable in the presence of transmission errors.

The predictor uses contributions from several samples in the present and previous fields. The weighting coefficients were chosen to minimize the mean-square prediction error measured over a wide range of picture material[2]. The coefficients used in this study are given in one column of Table 1.

A3b: Three-field multi-element predictor

This predictor was designed in the same manner as A3a above but elements from the second previous field have also been included. The coefficients are given in the final column of Table 1.

2.2 Category B: Adaptive prediction with transmission of the adaptation control signal

In these strategies, a choice is made between two or more linear predictors on the basis of the prediction error signal resulting from each. The choice of prediction is then signalled to the decoder. The first two systems chosen for investigation in this category switch between two predictors: the two-dimensional predictor, A1, and the pure interframe predictor, A2, given above. In stationary areas of the picture A2 gives the better prediction while in moving areas A1 probably gives the better prediction. The two systems investigated differ in the strategies used to select between the two predictors.

Table 1
Coefficients of multi-element fixed predictors

Coordinates as described in text			Predictor coefficients	
x	y	t	Predictor A3a	Predictor A3b
1	0	0	1.000	1.000
2	0	0	−0.471	−0.437
3	0	0	0.295	0.235
0	2	0	0.118	0.327
1	2	0		−0.195
0	−1	0	0.404	0.201
1	−1	0	−0.349	−0.158
2	−1	0	0.138	
3	−1	0	−0.104	
0	1	0	0.319	0.172
1	1	0	−0.394	−0.172
2	1	0	0.159	
3	1	0	−0.116	
0	0	1		0.460
1	0	1		0.497
2	0	1		0.277
3	0	1		−0.149
0	2	1		−0.291
1	2	1		0.198

B1: Sample-by-sample switched prediction

This strategy was proposed by FTZ of West Germany for a 34 Mbit/s *YUV* codec that uses DPCM coding together with Huffman coding of the quantized prediction error[3]. The basis of the strategy is to use, for each sample in turn, the predictor that yields the smaller prediction error. As it stands, such a strategy would require an unacceptably large signalling overhead. This overhead is greatly reduced by the use of hysteresis, which is introduced into the selection algorithm by weighting the current decision with contributions from previous decisions. Details of the decision weighting coefficients and of the coding scheme for the predictor signalling are given in Ref. 3.

B2a: Block-by-block switched prediction
(2 predictors)

This strategy has been proposed by Swedish Telecom for a 34 Mbit/s *YUV* codec[4]. In this codec, variable-length coding of the prediction error signal is also proposed. The signal is divided into blocks of 9 consecutive samples. The predictor chosen for a given block is the one which gives the smaller number of transmitted bits for that block. However, in our investigations the predictor that gives the smaller mean square prediction error averaged over a block is selected for that block. This alternative selection procedure has been used to simplify later comparison with other DPCM sytems. It is thought to give only slightly worse results than the original criterion. One bit of signalling information is required for each block. The signalling overhead is therefore about 0.11 bits per sample.

B2b: Block-by-block switched prediction
(3 predictors) with resetting

The third strategy in this category is similar to system B2a but uses three predictors, a simple two-dimensional predictor similar to A1, a pure interframe predictor and an interfield predictor with five elements, $g(1,0,0) = 1$, $g(0,1,0) = g(0,-1,0) = 0.5$, $g(1,1,0) = g(1,-1,0) = -0.5$. An additional feature is the inclusion of a resetting strategy which ensures that transmission errors do not propagate over more than 8 fields. The strategy was included in our investigation because it represented a later stage in the development of proposals for international standardization.

2.3 Category C: Adaptive prediction without explicit transmission of the adaptation control signal

In these strategies, the choice between predictors is made on the basis of information available at both the coder and the decoder, so that there is no signalling overhead. Three systems in this category were compared.

C1: Median Adaptive Prediction

This method is used in a *YUV* codec design by KDD of Japan[5]. An odd number of different predictions are generated and the prediction that has the median prediction value is selected (i.e. the value in the middle of an ordered list of prediction values). The decoder is able to make the same selection as the coder without requiring additional signalling. In our case, three predictors (A1, A2 and A3a) were used, so that the strategy ensures that either the best or the second best predictor is used for each sample. (The predictors used in our simulations are not the same as the three predictors described in Ref. 5. The choice of A1, A2 and A3a was made in order to test the median adaptive selection strategy in comparison with non-adaptive strategies.)

C2: Three-plane nonlinear predictor

This predictor was devised by Dukhovich and O'Neal[6]. The prediction is a weighted sum of three 'representatives' each of which is itself a switched adaptive, or nonlinear, estimate of the current sample based on previous samples. Each estimate uses samples in one of three mutually perpendicular planes in three-dimensional (horizontal, vertical, temporal) space. The details are given in Ref. 6.

C3: Spatial adaptive predictor

This simple strategy uses one of two predictors, the previous-sample predictor $g(1,0,0) = 1$ and the two-dimensional predictor $g(1,0,0) = 0.5$, $g(0,2,0) = 0.5$. For each sample, the second predictor is used unless the difference between the locally decoded signals corresponding to elements $g(1,0,0)$ and $g(1,2,0)$ is greater than a fixed threshold (set to 15 in our investigations), indicating a horizontal edge, in which case the first predictor is used. The strategy is published in Ref. 7 as the fallback predictor in a complex edge-adaptive algorithm which, in our investigations, gave poorer results than the simple fallback strategy used alone.

2.4 Category D: Continuously adaptive prediction

These strategies involve making small adjustments to the predictor on a sample-by-sample basis. Thus, the predictor adapts to compensate for movement in the picture or changes in the picture statistics. Two such systems which have been reported in the literature have been investigated.

D1: Pel-recursive motion-compensated predictor

This is a motion-compensated single-element interframe predictor and is used in a codec design of Bell Northern Labs[8]. The motion vector is updated at each sample by an amount which is proportional to the previous prediction error and to the local spatial gradient. Two similar versions of this prediction strategy have been investigated. The first, D1a, uses the full values of the gradient and the prediction error whereas the second, D1b, approximates these values by their signs (for instrumental simplicity). In both cases the magnitude of the correction to the motion vector that can be made at each sample is limited by multiplying the correction estimate by a small 'convergence parameter'. Several values of the convergence parameter were tried and an optimum value was found.

The updating of the motion vector is carried out independently within (16×8)-sample spatial blocks. At the start of each block in a frame, the initial motion vector is set equal to the last vector obtained for that block in the previous frame.

D2: LMS adaptive prediction

This technique is described by Alexander and Rajala[9]. It attempts to correct coefficients in the predictor that made a large contribution to a bad prediction. The coefficients are continuously corrected, sample by sample, by adding to each one the product of a convergence parameter, the prediction error at the previous sample and the value of the decoded signal that was used by the coefficient concerned. It can be shown that the resulting predictor approaches the minimum-mean-square predictor for the local picture statistics. Since the predictor corrections are calculated using previous picture information the decoder is able to keep in step with the coder without the transmission of additional information (in the absence of transmission errors).

Several variations of this LMS prediction technique were examined[10] and the results of two variations are presented here. In preliminary investigations, the LMS algorithm was used to adapt the coefficients of predictor A3a given in Table 1. Several values of the convergence value were tried and an optimum value was used for the results (labelled D2a) quoted in Section 3. However, later tests showed that the decoder was highly unstable in the presence of transmission errors. A more stable strategy was developed[10] in which the LMS technique was used to derive a stable predictor for each field. The coefficients of this predictor (labelled D2b in Section 3) are then transmitted to the decoder at the beginning of each field.

3. COMPARISON OF PREDICTION STRATEGIES

DPCM coders using the prediction strategies described above were simulated on a computer-based image processor and the results displayed on a sequence store. However, it was very difficult to make a clear comparison of the effectiveness of the different prediction strategies by subjective assessment of the picture quality. Therefore other methods, described in section 3.2, have been used to obtain an objective comparison of the different strategies.

3.1 Source pictures

Two sequences of moving pictures were used in the simulations. The first was the CCETT 'Voiture' sequence. This sequence contains some high-contrast detail with a variety of speeds and directions of motion. However, as a source sequence it is also noisy. As a contrast, a second sequence was used which consisted of a highly detailed portion of the 'Kiel Harbour' picture. This portion was moved horizontally at a speed of 2 pixels per field. The effect of camera integration at that speed was included in the sequence. Both these sequences can be considered as 'critical' test material.

Sixteen-frame, quarter-size, luminance-only sequences were used in the investigations. The pictures were sampled at 13.5 MHz and stored to 8-bit accuracy with black signal level to peak white signal level occupying a range of 224 levels. The first frame of each sequence is shown in Fig. 2.

The performance of some of the prediction strategies were also tested with a stationary picture. A sequence consisting of sixteen repeated frames of the central quarter of the EBU test picture 'Boy with Toys' was used for this purpose.

3.2 Techniques for comparison

At present, there is no generally accepted, universal measure of picture quality that takes into account the many different types of impairment that can be introduced by different coding systems. However, for all DPCM systems the nature of the impairments introduced are similar. Also, these impairments are mainly noise-like for broadcast-quality systems. Therefore, an acceptable comparative measure of picture quality for different DPCM systems is the signal-to-quantizing noise ratio (*S/N*).

In a DPCM system, the parameters of the non-linear quantizer determine the amount of quantization distortion or noise in the decoded picture. The output spacings of the inner quantizing levels, the mid-range levels and the outer range of the quantizer determine

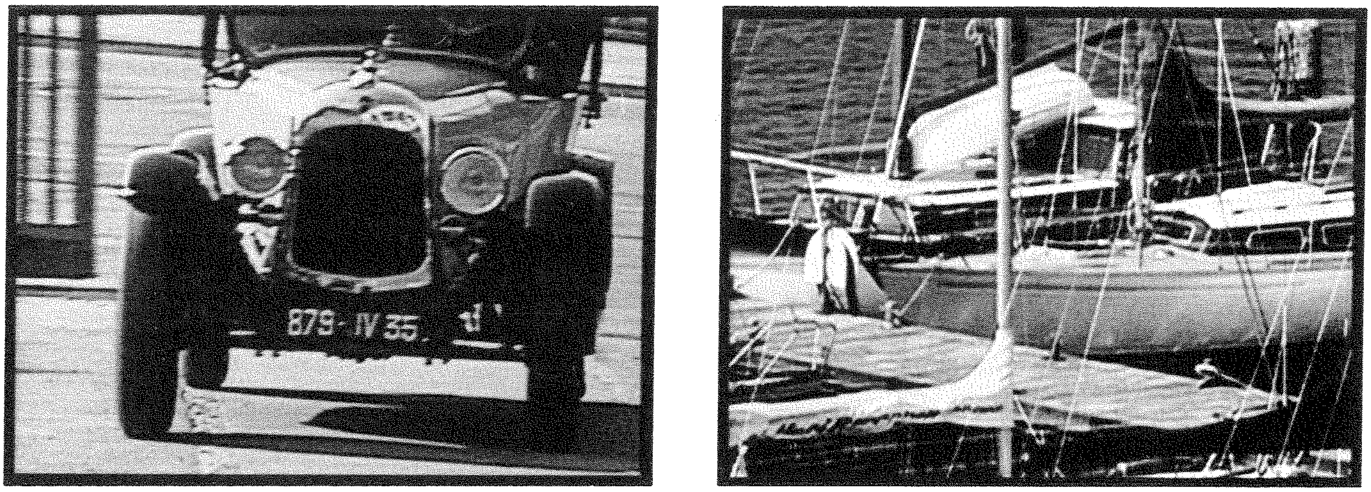


Fig. 2 - The first frames of the test sequences 'Voiture' and 'Kiel Harbour'.

the amount of granular noise, edge busyness and slope overload respectively. The influence of the predictor on output quality is mainly through the restrictions that it imposes on the spacing of levels in the quantizer. The form of those restrictions depends on whether or not the system involves variable-length coding.

In a system with no variable-length coding, the fixed number of output levels available for coding each sample must be arranged to cover almost the complete range of prediction error values. The best-quality coding in such a system, therefore, is probably obtained using a predictor for which the range of the prediction error is least. The range of the prediction error can be defined as the value not exceeded by 95% of all prediction errors. This assumes that a small degree of slope overload is acceptable. The 95% cumulative probability figure can be calculated from the probability density function (p.d.f.) of the unquantized prediction error and is a good comparative measure of the effectiveness of each prediction strategy for sytems not using variable-length coding.

When variable-length coding is used, the 95% figure is not necessarily a good measure because an occasional large (quantized) prediction error can be accommodated by the variable-length code and does not add significantly to the average bit rate. In this case the entropy of the quantized prediction error is a better guide to predictor performance. By suitable design of variable-length codes, the transmission bit rate per sample can be made to approach the entropy of the quantized prediction error signal.

When variable-length coding is used, what we really want to compare is the picture quality at a certain average bit rate, for example 3 bits/sample. As discussed earlier, when the impairments introduced by different systems are similar, then an acceptable measure of picture quality is the signal-to-quantizing noise ratio (*S/N*). Therefore, in order to compare the different prediction strategies, the *S/N* produced by a DPCM coder using a given prediction strategy is measured for a nonlinear quantizer which has been adjusted in each case to give an average entropy of 3 bits/sample. The best prediction algorithm minimizes both the entropy and the average quantizing noise power resulting from the use of a given nonlinear quantizer. The quantization noise can be further reduced by reducing the spacing of the quantizer output levels but this increases the entropy figure.

The comparative *S/N* figures at 3 bits/sample were obtained in the following manner for each prediction strategy:

(i) The *S/N* resulting from a typical non-linear quantizer that gives an entropy near 3 bits/sample was calculated from the difference between the coded and input pictures. The quantizer used is symmetrical, has 19 levels and is typical of those used in variable-length coding systems which code at around 3 bits/sample[3]. The output levels of this quantizer are listed in Table 2.

Table 2
19-level nonlinear quantizer used in the simulation

Input levels	Output level
0 to 2	0
3 to 7	5
8 to 13	10
14 to 20	17
21 to 28	24
29 to 38	33
39 to 50	44
51 to 64	57
65 to 79	72
80 to 255	87

(ii) The entropy of the quantized prediction error was measured.

(iii) An estimate was then made of the *S/N* that would result from a quantizer whose output levels had been adjusted to give an entropy of exactly 3 bits/sample. The Appendix shows that stretching or shrinking the quantizer causes the *S/N* to change by approximately 6 dB per bit of the entropy figure. A suitable correction can therefore easily be made to the *S/N* figures previously obtained for the typical quantizer.

3.3 Results of comparisons for moving picture sequences

Fig. 3 shows the comparative results using the two measures described above for the two moving picture sequences. The first measure is the range of the quantizer required for each predictor. The smaller the range, the better is the predictor. The second measure is the *S/N* at 3 bits/sample. For each measure, the lines drawn between the results are only intended to indicate that the measurements were obtained using a common source picture sequence. As a comparison, *S/N* results from a two-dimensional coding scheme based on the discrete cosine transform are also given.

Some features of interest that can be drawn from the results are as follows:

(i) The performance of many of the strategies are approximately similar.

(ii) The rank order of the results is similar for both measurement criteria. (Note that strategy B1, the one major exception to this result, is not really suitable for use in a fixed-length coding system because the signalling bit rate is itself variable).

(iii) The continuously adaptive LMS algorithm (D2a) is by far the best prediction strategy according to both criteria and on both sequences. The other continuously adaptive prediction strategy (D1a,b) based on motion estimation is not so effective. This is because in highly detailed areas of the picture the rapidly changing gradients can easily cause the motion estimate to diverge from the real value. It is thought that the algorithm may perform better if the motion estimate is derived from spatially low-pass filtered pictures.

(iv) Three-dimensional multi-element fixed predictors such as A3a and A3b compare favourably with the adaptive strategies.

(v) Of the switched prediction strategies involving explicit signalling (Category B), the one using sample-by-sample switching (B1) gives the worst *S/N* performance. This is because the signalling overhead for this strategy can become large for noisy or highly detailed pictures. For example, this overhead reached 0.62 bits/sample for the 'Kiel Harbour' sequence.

(vi) The two LMS-based schemes are the only ones that give a better average performance than two-dimensional transform coding. However, the signal-to-noise ratio is not a good measure for comparing the performance of transform coding and DPCM systems. With transform coding the noise introduced tends to be more visible because it is not concentrated only at edges, as with DPCM, but can appear as blocks of noise spreading from edges into surrounding plain areas.

Although these results have been obtained from a limited range of picture material, it is believed that they give a good indication of comparative performance on critical picture material.

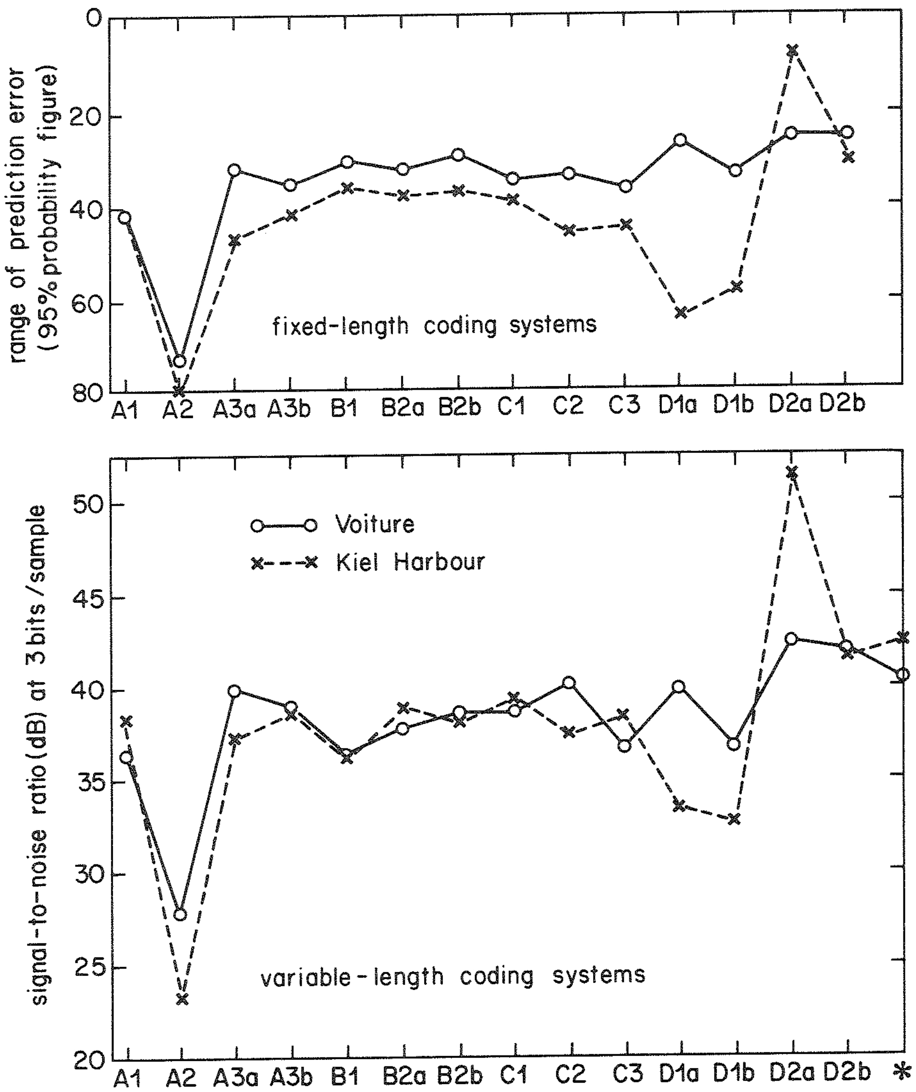
3.4 Results of comparisons for stationary picture material

The above results may be considered as 'worst case' for the various prediction strategies because of the critical nature of the test sequences. It is important that no strategy should fail drastically (as did the interframe predictor A2) and coding algorithms are usually assessed by their performance on critical sequences. However, typical picture material is less critical and contains both stationary and moving picture areas. The performance of the algorithms on stationary pictures is also a factor, therefore, which determines the performance for more typical picture material.

Comparative results for a selection of prediction algorithms with the (stationary) EBU test picture 'Boy with Toys' are shown in Fig. 4. Features of note are as follows:

(i) Again, the rank order of the results is similar for both measurement criteria.

(ii) As expected, the performance of the interframe predictor is excellent for



A1 – Typical two-dimensional predictor
A2 – Pure interframe predictor
A3a – Two-field multi-element predictor
A3b – Three-field multi-element predictor
B1 – Sample-by-sample switched prediction
B2a – Block-by-block switched prediction (2 predictors)
B2b – Block-by-block switched prediction (3 predictors) with resetting
C1 – Median adaptive prediction
C2 – Three-plane nonlinear predictor
C3 – Spatial adaptive predictor
D1a – Pel-recursive motion compensated predictor — gradient
D1b – Pel-recursive motion compensated predictor — sign only
D2a – LMS adaptive prediction — unstable
D2b – LMS adaptive prediction — stable
* – Two-dimensional DCT coding

Fig. 3 - Comparison of coding performance for moving picture sequences.

stationary pictures compared with its very poor performance for moving sequences. (The performance is not perfect because of quantization error feedback around the prediction loop.)

(iii) For stationary pictures, the performance of spatial predictors is poor compared with adaptive and non-adaptive predictors which use previous frame or field information.

(iv) On this picture the performance of the 2-dimensional DCT is similar to the performance of the spatial predictors.

Fig. 4 also shows the comparative result for a multi-element, two-line, spatial predictor which was designed to minimize the mean square prediction error over a range of picture material. This predictor has nine elements and is listed in Table 3. The performance is better than that of the other purely spatial predictors, A1 and C3, but is not as good as the multi-element predictors A3a and A3b. This 9-element spatial predictor is similar to one used in DPCM equipment which was constructed by BBC Research Department to code 4:2:2 *YUV* signals into 140 Mbit/s[1]. At this bit rate, 6 bits/sample are available for coding each of the luminance and chrominance components and very good coded picture quality is obtained using this predictor.

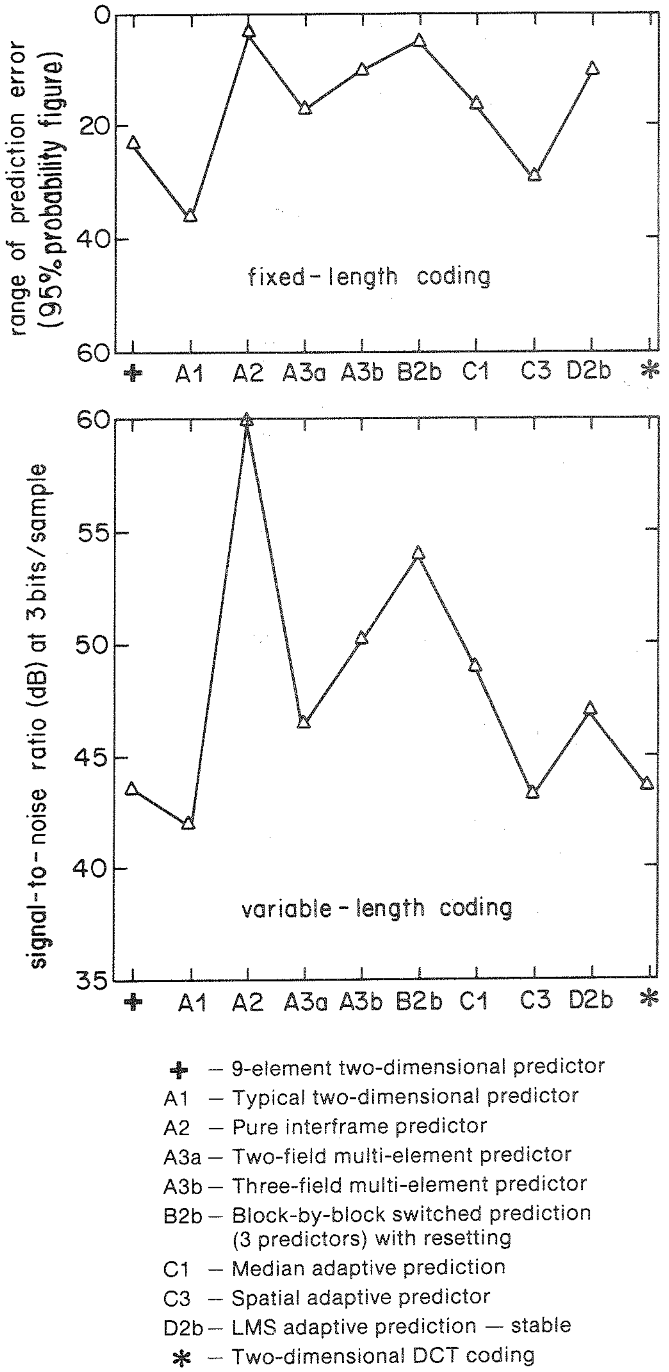


Fig. 4 - Comparison of coding performance for stationary picture.

Table 3
Coefficients of 9-element two-dimensional predictor

Coordinates as described in text x	y	t	Predictor coefficients
1	0	0	1.098
2	0	0	−0.612
3	0	0	0.497
4	0	0	−0.297
5	0	0	0.154
6	0	0	−0.021
−1	2	0	0.107
0	2	0	0.360
1	2	0	−0.315

3.5 Comparison between fixed- and variable-length coding

A detailed assessment of the relative worth of fixed-length and variable-length coding for each of the prediction strategies is not presented here. However, a few measurements were carried out, using a typical quantizer suitable for a fixed bit rate of about 4 bits per sample. Such a quantizer gave noise power figures similar to those calculated above for a 3-bit variable-length coding system for a wide range of prediction strategies. This indicates that the benefit of variable-length coding is about 1 bit per sample or 6 dB, for any system.

4. MEASUREMENTS OF ERROR PERFORMANCE

A further basis for comparison between prediction strategies is their effect on the decoded picture quality in the presence of transmission errors. The design of variable-length codes[11], quantizers and error-protection strategies also affects this behaviour; nevertheless, the need for too much complexity and bit-rate penalty in those areas can be avoided if the prediction strategy itself is stable.

An attempt was made at measuring this stability for some of the prediction strategies by simulating the operation of a coder and decoder (without quantization) in cascade, with errors inserted into the prediction error signal at the input to the decoder. Two different kinds of error were investigated. For the first kind, an average of 1 in 10^4 samples of the prediction error signal were chosen on a pseudo-random basis (but the same samples for each prediction strategy) and replaced with values of ±128, choosing the worse of the two possibilities in each case. This simulates well the effect of errors in a fixed-length coding scheme and, in some circumstances, in a variable-length coding scheme. Other possible effects in a variable-length coding system can be simulated by a second kind of error, produced by replacing each line containing a pseudo-randomly chosen sample with the previous line in the field. However, it was found that the two kinds of error had similar effects on the performance of the prediction strategies and so, for simplicity, only results involving the first kind of error are presented.

In order to measure the effect of the errors, a difference signal between the input and the decoded signal with errors was generated. An informal subjective assessment indicated that, of three different quality measures based on the p.d.f. of this difference signal (mean square value, mean absolute value and proportion of absolute differences greater than 10), the mean square value provided as good a correlation with subjective impressions as either of the other measures.

Fig. 5 shows the results of a comparison between some of the strategies based on this measure, expressed as *S/N* values. Only the 'Voiture' sequence was used in this comparison.

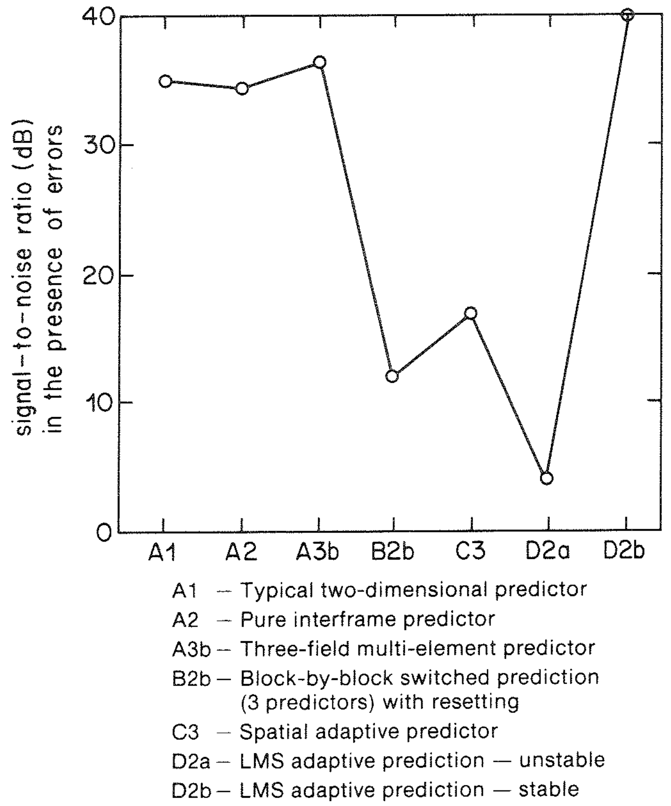


Fig. 5 - Comparison of error performance.

The results of this investigation can be summarized as follows:

(i) Fixed predictors, even complex ones (provided stability is taken into account in their design), lead to very good performance in the presence of transmission errors. Their use avoids the need for an explicit resetting strategy to ensure that the decoder is in step with the coder.

(ii) The adaptive systems chosen for investigation behave poorly in the presence of errors, even in the case of strategy B2b which uses resetting to limit error propagation.

(iii) The original version of the LMS algorithm has disastrous error performance. However, the version incorporating a stability criterion and transmission of the predictor coefficients is the most stable of all.

5. CONCLUSIONS

Measurements have been carried out to compare the performance of several DPCM prediction strategies that have been reported in the literature, with particular application to the coding of *YUV* signals at bit rates between 30 and 60 Mbit/s. The measurements are based on simulations of the various coding schemes for critical pictures and picture sequences.

The results show that, for critical sequences, the continuously adaptive LMS algorithm has the best performance but is highly unstable in the presence of transmission errors. The best stable predictor is the LMS-based predictor the coefficients of which are transmitted each field. Further discussion of the LMS technique is given in a companion Report[10].

The performance in terms of picture quality of several of the prediction strategies appears to be similar. Therefore, if a choice between them is to be made, other factors must be taken into account such as hardware complexity and coder/decoder performance in the presence of transmission errors.

In particular, the two- and three-field multi-element predictors are straightforward to implement and give very good performance in terms of coded picture quality and resistance to the effects of transmission errors.

6. ACKNOWLEDGEMENTS

The authors thank CCETT of France and the Deutsche Bundespost respectively for providing the 'Voiture' and 'Kiel Harbour' source pictures.

7. REFERENCES

1. WELLS, N.D. 1984. Digital video: *YUV* bit rate reduction for broadcasting applications. Tenth International Broadcasting Convention, Brighton, Sept. 21-25 1984. IEE Publication No. 240, pp 338-340.

2. MOORE, T.A. 1986. Applications of linear minimum mean square optimisation techniques in video coding. Proc. IEE Conference on Image Processing and its Applications, London, June 1986.

3. BULEY, H. and STENGER, L. 1984. Inter/intraframe coding of colour TV signals for transmission at the third level of the digital hierarchy. *Proc. IEEE,* 73, 4, April 1985.

4. WEISS, P., BENGTSSON, L. and CHRISTENSSON, B. 1985. Adaptive DPCM-coding of broadcast TV-signals at 34 Mbit/s. COST 211 bis contribution (unpublished).

5. MURAKAMI, H. et al. 1983. A 15 Mbit/s universal codec for TV signals using a median adaptive prediction coding method. Proc. 6th Intern. Conf. on Dig. Sat. Comm. Sept. 19-23 1983, Phoenix, Arizona, USA.

6. DUKHOVICH, I.A. and O'NEAL, Jr., J.B. 1977. A three-dimensional spatial non-linear predictor for television. *IEEE Trans. Comm.* **26**, 5, May 1978.

7. KRETZ, F. 1983. Edges in visual scenes and sequences: application to filtering, sampling and adaptive DPCM coding. In: Image sequence processing and dynamic scene analysis (ed. T.S. Huang), Springer-Verlag, Berlin and Heidelberg, 1983.

8. SABRI, S. 1983. Movement-compensated interframe prediction for NTSC colour TV signals. In: Image sequence processing and dynamic scene analysis (ed. T.S. Huang), Springer-Verlag, Berlin and Heidelberg, 1983.

9. ALEXANDER, S.T. and RAJALA, S.A. 1984. Image compression results using the LMS adaptive algorithm. *IEEE Trans. on ASSP*, **33**, 3, June 1985.

10. KNEE, M.J. and WELLS, N.D. 1988. LMS prediction for differential image coding. BBC Research Department Report in course of preparation.

11. KNEE, M.J. 1987. Designing variable-length codes with good error-recovery properties. BBC Research Department Report No. BBC RD 1987/19.

APPENDIX

Variation of *S/N* with Entropy

In this Appendix, we illustrate that if a certain predictor and nonlinear quantizer give a *S/N* of S_1 dB and a quantized prediction error entropy H_1 then a first-order approximation for S_2, the *S/N* resulting from stretching or shrinking the quantizer to give an entropy of H_2, is as follows:

$$S_2 = S_1 + 6(H_2 - H_1) \text{ dB}$$

First, we define for the quantizer a function $c(x)$, as follows:

$$c(x) = d(i+1) - d(i)$$

where x is the unquantized prediction error and $d(i)$ and $d(i+1)$ are adjacent output decision levels in the quantizer such that $d(i) < x \leq d(i+1)$. This function can be thought of as a measure of the 'coarseness' of the quantizer at prediction errors around x. The entropy H_1 is given by

$$H_1 = -\sum_i P(i)\log_2 [P(i)]$$

where $P(i) = \text{Prob}\{d(i) < x \leq d(i+1)\}$ is the p.d.f. of the quantized prediction error. $P(i)$ can also be expressed as follows:

$$P(i) = \sum_{x=d(i)}^{d(i+1)} p(x) \qquad \text{(where } p(x) \text{ is the p.d.f. of } x\text{)}$$

and, if we assume that $p(x)$ is approximately constant between any two adjacent decision levels,

$$P(i) = p(x)c(x)$$

for any x between $d(i)$ and $d(i+1)$.

Hence, we can approximate the entropy H_1 by

$$H_1 = -\sum_{x} p(x)\log_2 [p(x)c(x)]$$

The S/N S_1 is given by

$$S_1 = \text{k} - 10\log_{10} \sum_{i} \sum_{x=d(i)}^{d(i+1)} p(x)[x-q(i)]^2$$

where k is a constant and $q(i)$ is the reconstruction level lying between $d(i)$ and $d(i+1)$. This can be approximated by

$$S_1 = \text{k}' - 10\log_{10} \sum P(i)\, c(x)^3 \qquad \text{where } d(i) < x \leq d(i+1)$$

$$= \text{k}' - 10\log_{10} \sum p(x)\, c(x)^2$$

where k′ is another constant.

We now stretch (or shrink) the quantizer by multiplying all its decision and reconstruction levels by a constant, w. If w is near to 1, we may assume that $p(\text{w}x) = p(x)$ for all values of x. Thus, the effect of stretching the quantizer in the expressions derived above for entropy and S/N is simply to multiply all the values of $c(x)$ by w. We may therefore write expressions for the entropy H_2 and the S/N S_2 for the new quantizer as follows:

$$H_2 = -\sum_{x} p(x)\log_2 [p(x)c(x)\text{w}]$$

$$= H_1 - \log_2 \text{w}$$

$$S_2 = \text{k}' - 10\log_{10} \sum_{x} p(x)c(x)^2\text{w}^2$$

$$= S_1 - 20\log_{10} \text{w}$$

Hence

$$S_2 = S_1 + \frac{20(H_2 - H_1)}{\log_2 10}$$

$$\simeq S_1 + 6(H_2 - H_1)$$

as required.